Demography or selection on linked cultural traits or genes? Investigating the driver of low mtDNA diversity in the sperm whale using complementary mitochondrial and nuclear genome analyses

Morin, Phillip; Foote, Andrew; Baker, C. Scott; Hancock-Hanser, Brittany; Kaschner, Kristin; Mate, Bruce; Mesnick, Sarah; Pease, Victoria; Rosel, Patricia; Alexander, Alana

Molecular Ecology

DOI:
10.1111/mec.14698

Published: 01/06/2018

Peer reviewed version

Demography or selection on linked cultural traits or genes? Investigating the driver of low mtDNA diversity in the sperm whale using complementary mitochondrial and nuclear genome analyses

Phillip A. Morin[1]*, Andrew D. Foote[2], C. Scott Baker[3,4], Brittany L. Hancock-Hanser[1], Kristin Kaschner[5], Bruce R. Mate[3,4], Sarah L. Mesnick[1], Victoria L. Pease[1], Patricia E. Rosel[6], Alana Alexander[7,8]

[1]Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, California, USA

[2]Molecular Ecology and Fisheries Genetics Laboratory, School of Biological Sciences, Bangor University, Bangor, Gwynedd, LL57 2UW, UK

[3]Marine Mammal Institute, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Drive, Newport, OR 97365, USA,

[4]Department of Fisheries and Wildlife, College of Agricultural Sciences, 104 Nash Hall, Corvallis, OR 97330, USA

[5]Department of Biometry and Environmental System Analysis, Albert-Ludwigs-University of Freiburg, Tennenbacher Strasse 4, 79106 Freiburg, Germany

[6]Southeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Lafayette, Louisiana, USA

[7]Biodiversity Institute, University of Kansas, 1345 Jayhawk Blvd, Lawrence, KS 66045, USA

[8]Current address: Department of Anatomy, University of Otago, Dunedin 9016, New Zealand

*Corresponding author: phillip.morin@noaa.gov

Short title: Global phylogeography of the sperm whale

Keywords: cetacean, PSMC, population structure, Pleistocene

Abstract

Mitochondrial DNA has been heavily utilized in phylogeography studies for several decades. However, underlying patterns of demography and phylogeography may be misrepresented due to coalescence stochasticity, selection, variation in mutation rates, and cultural hitchhiking (linkage of genetic variation to culturally transmitted traits affecting fitness). Cultural hitchhiking has been suggested as an explanation for low genetic diversity in species with strong social structures, counteracting even high mobility, abundance and limited barriers to dispersal. One such species is the sperm whale, which shows very limited phylogeographic structure and low mtDNA diversity despite a worldwide distribution and large population. Here, we use analyses of 175 globally distributed mitogenomes and three nuclear genomes to evaluate hypotheses of a population bottleneck/expansion versus a selective sweep due to cultural-hitchhiking or selection on mtDNA as the mechanism contributing to low worldwide mitochondrial diversity in sperm whales. In contrast to mtDNA control region (CR) data, mitogenome haplotypes are largely ocean-specific, with only one of 80 shared between the Atlantic and Pacific. Demographic analyses of nuclear genomes suggest low mtDNA diversity is consistent with a global reduction in population size that ended approximately 125,000 years ago, correlated with the Eemian interglacial. Phylogeographic analysis suggests that extant sperm whales descend from maternal lineages endemic to the Pacific during the period of reduced abundance, and have subsequently colonized the Atlantic several times. Results highlight the apparent impact of past climate change, and suggest selection and hitchhiking are not the sole processes responsible for low mtDNA diversity in this highly social species.

Introduction

There is a growing body of literature indicating that the commonly held tenet of population genetics, that genetic diversity should be correlated with population size, is often violated, particularly for mitochondrial DNA (mtDNA) (Bazin *et al.* 2006). When mtDNA diversity is found to be low in currently abundant populations, several hypotheses have been invoked to explain the discord, including population bottlenecks (e.g., Hoelzel *et al.* 2002; O'Brien 1994), selection on mtDNA (e.g., Finch *et al.* 2014; Foote *et al.* 2011), variation in mutation rates (Lyrholm *et al.* 1996), and cultural hitchhiking (linkage of genetic variation to culturally transmitted traits affecting fitness; Kopps *et al.* 2014; Premo & Hublin 2009; Whitehead 1998, 2005).

Among globally distributed large whales, most baleen whales exhibit high mtDNA diversity relative to toothed whales and are divided into multiple subspecies and genetically distinct populations (e.g., fin, humpback, gray and blue whales; Archer *et al.* 2013; Baker *et al.* 2013; Jackson *et al.* 2014; Lang *et al.* 2014; Leduc *et al.* 2016). Among toothed whales, however, unusually low mitochondrial DNA (mtDNA) diversity in some of the social odontocetes (e.g., sperm, pilot, killer, and false-killer whales; Alexander *et al.* 2016; Alexander *et al.* 2013; Hoelzel *et al.* 2002; Martien *et al.* 2014; Van Cise *et al.* 2016) has limited power to infer population structure, phylogeography and historical demography using traditional genetic tools. The sperm whale (*Physeter macrocephalus*) is particularly enigmatic in this respect, as it is one of the most cosmopolitan and abundant of the large odontocetes, and known to move over large ranges of up to thousands of kilometers over annual or longer time periods (Mizroch & Rice 2013; Steiner *et al.* 2012; Straley *et al.* 2014), yet it exhibits low mtDNA diversity and evidence of female philopatry (Alexander *et al.* 2016; Lyrholm & Gyllensten 1998; Lyrholm *et al.* 1999; Mesnick *et al.* 2011).

In addition to the sperm whale's broad distribution throughout the world's oceans, it is considered relatively abundant even post-industrial whaling, with the global population size estimated in 2002 at 360,000, approximately 1/3 of its pre-whaling population size (Whitehead 2002). Previous genetic studies based on mtDNA control region (CR)

sequences showed haplotype frequency differences among oceans, major within-ocean geographic regions and female-led social groups, but limited phylogeographic signal, largely due to the low CR diversity and occurrence of three common haplotypes in all large ocean basins (Alexander *et al.* 2016; Engelhaupt *et al.* 2009; Lyrholm & Gyllensten 1998; Mesnick *et al.* 2011; Whitehead 1998).

In other species with low levels of CR diversity, complete mitochondrial genome (mitogenome) sequences have been used to detect phylogeographic structure and estimate divergence times and historical demography with higher precision (e.g., Archer *et al.* 2013; Buddhakosai *et al.* 2016; Morin *et al.* 2010; Morin *et al.* 2015; Shamblin *et al.* 2012). Several hypotheses have been proposed to explain the extremely low mtDNA diversity in sperm whales, including a population bottleneck (Lyrholm & Gyllensten 1998; Lyrholm *et al.* 1996), low mutation rate (Lyrholm *et al.* 1996; Whitehead 1998), stochastic variation in maternal lineage survival (Amos 1999; Tiedemann & Milinkovitch 1999), cultural hitchhiking (Whitehead 1998, 2005; Whitehead *et al.* 2017), and selective constraints on CR sequences that limit accumulation of variation and lead to saturation of sites free to vary (Alexander *et al.* 2013). Alexander et al. (2013) used complete mitogenome sequences (N=17) to evaluate support for some of these hypotheses, and concluded that the data were consistent with a bottleneck or selective sweep (involving selection on mitochondrial protein-coding regions or cultural hitchhiking), and found no evidence for selective constraint on the control region or slow substitution rates in sperm whale mtDNA relative to other cetaceans. In this study, we analyze complete mitogenomes from 175 globally distributed sperm whale samples to investigate phylogeography, selection, and historical demography. We compare demographic inferences from mitogenomes and nuclear genomes to evaluate the previously proposed hypotheses of a population bottleneck/expansion versus a selective sweep due to cultural-hitchhiking or selection on mtDNA as the mechanism contributing to low worldwide mitochondrial diversity in sperm whales.

Materials and Methods

DNA extraction and library preparation

Sperm whale tissue samples (n=158) collected by live-biopsy or from dead stranded animals were stored in salt-saturated 20% DMSO at -20°C in the US National Marine Fisheries Service (NMFS) Marine Mammal and Sea Turtle Research (MMASTR) Collection at the Southwest Fisheries Science Center (SWFSC), or in 70% ethanol at -20°C at the Oregon State University Cetacean Conservation and Genomics Laboratory (CCGL). Sample information is in supplemental Table S1. DNA from SWFSC samples was extracted from tissue samples using either a silica-membrane method (Qiaxtractor® DX reagents, Qiagen, Valencia, CA, USA) or a simple salt-precipitation procedure (Miller *et al.* 1988). DNA libraries for these samples were constructed and enriched for mitochondrial DNA according to Hancock-Hanser et al. (2013), and sequenced in 3 pools of 49-66 samples with Illumina GEII, HiSeq (100 bp) and NextSeq (75 bp) single-end reads. Eleven samples were repeated in two pools to increase read depth of coverage. CCGL samples were extracted using a standard phenol/chloroform method (Sambrook *et al.* 1989), modified for smaller samples (Baker *et al.* 1994). Libraries for these samples were constructed from long-range PCR products following Alexander et al. (2013), individually-barcoded and prepared for sequencing using a Nextera XT DNA Sample Preparation Kit (Illumina, La Jolla, CA, USA). Sample libraries were pooled and sequenced in three Illumina MiSeq paired-end runs (two 250 bp, one 75 bp).

Sequence read quality control and assembly

Assembly of sequence reads to the reference mitogenome (KC312603) was performed using custom scripts in R (R Core Team 2014) and publicly available programs as previously described (Hancock-Hanser et al. 2013; Dryad data repository doi: 10.5061/dryad.cv35b). The reference mitogenome was modified to improve assembly coverage at the 'ends' of the linearized mitogenome by adding 40 bp from each end of the sequence to the opposite end (so that reads could map across the artificial break point of the linearized sequence). Nucleotide sites in the consensus sequence for each sample were called "N" if there were <3 reads, <9 reads where there was nucleotide variation at the site among reads, or >9 reads where a single nucleotide did not represent >70% of the reads. All sequences were aligned and visually inspected in the program GENEIOUS (V. 6.0.5, Biomatters, Auckland, New Zealand), with indels and unique variants identified in

GENEIOUS then verified by visual inspection of the read alignments of individual sample assemblies in the BAM files. New and 17 previously published sequences (Alexander *et al.* 2013) were aligned and checked for frame-shift indels and coding sequence start and stop codons within protein-coding regions, based on published annotation of the reference sperm whale mitogenome.

Genetic diversity analyses and diagnosability

The R package strataG (Archer *et al.* 2017a) was used to calculate haplotype and nucleotide diversity by geographically-defined population (stratum). To test if haplotypic diversity and nucleotide diversity within strata were significantly different from one another, we estimated the variance of each measure via a stratified bootstrap (supplemental methods). In each iteration of the bootstrap, individuals were randomly chosen with replacement from each stratum, with sample sizes determined by the empirical data set. At each of the 1000 bootstrap replicates, differences in nucleotide and haplotype diversity were calculated between each stratum (stratum 1 diversity - stratum 2 diversity, for each measurement). Strata were considered to have significantly different haplotype and/or nucleotide diversity if the observed difference was in the lower 5% of the distribution (strata 2 significantly greater diversity than stratum 1), or in the upper 5% of the distribution (strata 1 significantly greater than stratum 2).

Diagnosability is a measure of the ability to correctly determine whether a specimen of unknown origin can be correctly assigned to a group based on a trait or traits (Archer *et al.* 2017b). To test the diagnosability of mitogenome sequences between Atlantic and Pacific sperm whales, we conducted a Random Forest analysis (Breiman 2001) to create a model to classify samples to ocean basins following Archer et al. (2017b). The Random Forest was built with 50,000 trees. Each tree was built using a random draw of 25 samples (half of the smallest oceanic sample size of $n = 50$ for the Atlantic, including the Gulf of Mexico and Mediterranean Sea) from each ocean basin to avoid biasing the model due to differences in sample size. The upper and lower confidence intervals of the correct classification score as well as the expected classification score (prior) were calculated as described in Archer et al. (2017b).

186

187 *Phylogenetic analyses*

188 Time-calibrated phylogenetic analysis was performed using two methods: creating a prior
189 distribution for the time to most recent ancestor (TMRCA) using a two-phase process
190 (described below), and estimating the TMRCA using previously published substitution
191 rates (see *Demographic reconstruction* section). The two-phase method utilized BEAST
192 (v 1.8; Drummond *et al.* 2012). Aligned coding sequences were extracted based on
193 published start/stop codons of the reference sequence. Optimal substitution models for
194 individual loci were determined based on Bayesian Information Content (BIC) using
195 PartitionFinder (v 1.1.1; Lanfear *et al.* 2012). All loci were constrained to having the
196 same underlying topology and clock rate (as loci on the mtDNA are fully linked without
197 recombination). Phase 1 estimated the TMRCA for all sperm whale mitogenomes based
198 on the coding partitions from four sperm whale haplotypes (mtGen11, 22, 30, 47)
199 selected to represent the four major clades identified in a maximum likelihood tree of all
200 full mitogenome unique haplotypes (PhyML v. 2.2.0; Guindon *et al.* 2009, implemented
201 in Geneious), along with aligned coding region sequences from the pygmy sperm whale
202 (*Kogia breviceps*, accession No. NC005272) and four beaked whale species (*Berardius*
203 *bairdii*, NC005274; *Ziphius cavirostris*, KC776696; *Mesoplodon densirostris*,
204 KF032860; *Hyperodon ampulatus*, NC005273). The Bayesian phylogeny was calibrated
205 using lognormal distributions on two calibration nodes: Odontoceti (34.67 Myr: 95% CI
206 29.93 – 40.17 Myr) (McGowen *et al.* 2009) and Ziphiidae (22 Myr: 14.86 – 32.56 Myr)
207 (Dornburg *et al.* 2012; McGowen *et al.* 2009). The ucld prior was set to a uniform
208 distribution (min. 1E-6, max. 1), the tree prior was speciation, Yule process, and the
209 chain was 100M MCMC steps, logged every 10k steps. Convergence of two replicates
210 and mixing were checked using TRACER (v1.5; Rambaut *et al.* 2014) and RWTY
211 (Warren *et al.* 2017a). The maximum clade credibility tree was generated with
212 TreeAnnotator (v1.8.1) in the BEAST software cluster (Drummond *et al.* 2012) after
213 removal of the first 10% of trees.

214

215 The phylogenetic analysis of all sperm whale mitogenomes (Phase II) was conducted as
216 above, separately for all complete mitogenome unique haplotypes (N=80) and unique

haplotypes of concatenated coding loci (N=60), with TMRCA calibration based on the lognormal distribution (log(stdev = 0.2)) of the median TMRCA estimate of sperm whales from Phase I (136.7 thousand years ago (KYA), 95% CI 85.2 – 201.1 KYA) following previously described methods (Morin *et al.* 2010; Morin *et al.* 2015). A strict clock was used with a constant size coalescent tree model. 100M MCMC steps and a single mutation model (TN93), selected based on the AIC in jModeltest v2.1 (Darriba *et al.* 2012), were used for the full-length mitogenome sequences. For the concatenated protein coding loci, we used 10M MCMC steps and 3rd position sites only, with two mutation models applied to *NADH6* and all other genes combined (HKY and TrN, respectively), based on analysis of individual loci in PartitionFinder. For both analyses, convergence was checked based on four replicate runs using TRACER v1.5 and RWTY.

A haplotype median joining network (MJN: Bandelt *et al.* 1999) was created using the program POPART (Leigh & Bryant 2015) with default settings.

Demographic reconstruction

We used the codon-partitioned concatenated protein-coding regions of the sperm whale mitogenomes (excluding *NADH6* following Alexander *et al.* 2013; Ho & Lanfear 2010) in a skygrid analysis (implemented in BEAST v1.8.3; Drummond *et al.* 2012) and a skyline analysis (implemented in BEAST v2.3.0; Bouckaert *et al.* 2014). All 175 samples were included in both analyses to approximate a "balanced" sampling strategy (multiple samples from multiple populations) found to result in the least bias in inferring demographic change (Heller *et al.* 2013), and a separate GTR+G+I substitution model was used for each codon partition. Prior shapes for parameters within the substitution models and the relative rates for each codon position were derived from Alexander et al. (2013) (summarized in supplemental Table S2). For the skygrid analysis, two chains of 100M states, sampling every 100,000 states, were run. For the skyline analysis, two chains of 50 million states with sampling every 10,000 states were carried out. Convergence was assessed for the two analyses by comparing posterior distributions between chains in TRACER v1.6, and for topologies through RWTY. Demographic

reconstructions were generated through TRACER. Model comparison was performed in TRACER using AICM with 1,000 bootstraps (Baele *et al.* 2012).

For demographic analysis based on the nuclear genome, we employed pairwise sequentially Markovian coalescent (PSMC; Li & Durbin 2011) analysis of three sperm whale high-coverage genomes to infer changes in effective population size (*Ne*). We aligned published short read data from three sperm whales (GMX-SRS38925; BioSample SAMN01906698 (reference genome, Gulf of Mexico); read files SRR680161, SRR680169, SRR674482; SC991024-177-1; BioSample SAMN06187412 (Pacific); read files SRR5136496, SRR5136506, SRR5136508, SRR5146847; BioSample SAMN06187413 (Indian); read files SRR5136491, SRR5136493, SRR5136497, SRR5146843), to the repeat-masked sperm whale reference genome v2.0.2 (accession No. GCA_000472045.1; Warren *et al.* 2017b). Details of read quality filtering and assembly are provided with supplemental Figure S1. The PSMC plot was scaled to an autosomal mutation rate (μ_A) of 2.9 x 10^{-8} substitutions per nucleotide per generation (Dornburg *et al.* 2012) and a generation time of 31.9 years (Taylor *et al.* 2007), and we conducted 100 bootstrap resamplings on all PSMC analyses. Sensitivity testing for different mutation rates and generation times were conducted by re-scaling the PSMC plots (see supplemental Figure S1). While PSMC can provide useful insights into demographic change inferred from changes in coalescent rates (Li & Durbin 2011), it is also sensitive to changes in population structure (Mazet *et al.* 2016). Therefore, we extended the PSMC analyses presented by Warren et al. (2017b) by also carrying out PSMC analysis of a pseudo-diploid genome made by randomly sampling an allele at each site from each of the individual genome assemblies using seqtk (https://github.com/lh3/seqtk; see supplemental Figure S1 for details). The pseudo-diploid analysis provides information on changes in the rate of coalescence between two individual genomes through time, and therefore the timing of changes in population structure relative to the changes in *Ne* inferred by the single genome PSMC analyses (see Cahill *et al.* 2016; Chikhi *et al.* 2018).

Ancestral range reconstruction

For biogeographic ancestral range reconstruction analyses, the full mitogenome phylogeny was pruned to twelve clades that had high posterior support. A custom R-script was then used to prune the tree topology to one branch per clade (N. Matzke; http://phylo.wikidot.com/example-biogeobears-scripts#pruning_a_tree). The terminal branches were assigned to ocean basin(s) based on the origins of samples found within the non-pruned clades, and the tree and geographic information were used for biogeographic analysis using the R package BioGeoBEARS (Matzke 2012, 2013). The package was used to compare three basic models (DEC, DIVALIKE, BAYARELIKE), and two forms of each model (with and without founder-event cladogenesis; Matzke 2014). These models allowed for different combinations of range expansion, range splitting (vicariance), and extinction, with the model conferring the highest likelihood on the tip data selected based on AIC. Likelihood ratio tests were used to compare the nested models, with and without founder-event cladogenesis. Clades were subsampled to test for the effect of uneven sampling (see methods in the legend of supplemental Figure S2).

Suitable-habitat maps were generated with the AquaMaps approach to species distribution modelling (Kaschner *et al.* 2016; Kaschner *et al.* 2011; Ready *et al.* 2010) based on physical and oceanographic parameters for the last glacial maximum (LGM, ~20 KYA), current, and future conditions (year 2100, based on IPCC A2 emissions scenario). Since adult male and female sperm whales have different habitat use patterns, with males migrating to higher latitudes to feed, we altered the global sperm whale suitable habitat model (males and females) to generate female-specific models by setting the minimum mean sea surface temperature (minSST, preferred minSST) to approximate the observed summer distributions of adult female and mixed family groups in the North Pacific identified from commercial whaling records (Ivashchenko *et al.* 2014). The adjusted mean SST does not perfectly reflect the summer SST patterns, but allows approximate habitat models for the different time periods. All models exclude the primary production envelope because this information is not available for the LGM, but comparison of the present-day distribution maps with and without primary production indicated little effect of this parameter (supplemental Figure S3 A-D, www.aquamaps.org, supported by previous research that suggested little correlation of

primary production with sperm whale distributional patterns: Jaquet & Whitehead 1996). Total sizes of distributions were subsequently calculated for both sexes in different ocean basins for all three time periods. We calculated areas using both a probability threshold of 0.0 (approximating the mean annual native range of a species including some potentially lesser utilized habitat) and ≥ 0.6 (shown to correspond to the core habitat of a species; Kaschner *et al.* 2011).

Selection analyses on mitogenome data

We aligned the codon-partitioned concatenated protein coding regions of newly generated sperm whale mitogenomes with the multispecies mitogenome alignment of Alexander et al. (2013), and updated this with more recently published cetacean mitogenomes/refseq versions of cetacean mitogenomes (see supplemental Table S3 for all accession numbers). Using the priors detailed in supplemental Table S4, and a Yule model of speciation, we ran two chains of 100 M generations sampling every 1,000 states in BEAST v2.3.0. After checking for convergence using TRACER and AWTY (Nylander *et al.* 2008; Wilgenbusch *et al.* 2004), we used the tree we obtained for tree-based analyses of selection using TreeSAAP v3.2 (Woolley *et al.* 2003) and PAML v4.9 (Yang 2007). In addition to these methods, we estimated selection for each codon using HyPhy (Pond *et al.* 2005) as implemented in MEGA v6.06 (Tamura *et al.* 2013). We further examined selection using the MEME (Murrell *et al.* 2012) and FUBAR methods (Murrell *et al.* 2013), as implemented on the Datamonkey webserver (Delport *et al.* 2010; Pond & Frost 2005). MEME is designed to detect sites that have experienced episodic diversification, while FUBAR is designed to detect sites that have experienced pervasive diversification. We used both methods to detect any sites that have experienced positive selection in the sperm whale. We conducted all of these analyses using just one representative sperm whale haplotype (mtGen01), and GTR models of nucleotide substitution (or REV models where GTR was not available). For investigating specific sites inferred to be under positive selection we used $\alpha = 0.05$ as the threshold for statistical significance in HyPhy, MEME and FUBAR. Within PAML M8, we used Bayes Empirical Bayes (BEB) (Yang *et al.* 2005), with a threshold of $P > 95\%$. TreeSAAP results were restricted to putative sites under selection where at least one

property had a magnitude category of six or more in every pairwise comparison between the representative sperm whale and other cetaceans.

Following Caballero et al. (2015), 3D homology models of proteins for genes where positive selection was detected in sperm whales (*ND1, ND2, ATP8, COX3, ND4L, COX3, ND4, ND5, CYTB*) were generated by the SWISS-MODEL server (Schwede *et al.* 2003), using the templates specified in supplemental Figure S4 (best-fitting model based out of the top four templates for each gene region), and visualized using UCSF Chimaera v1.11 (Pettersen *et al.* 2004). This 3D model was annotated with domains based on alignment to the best-fitting template. We annotated a 2D model constructed with Protter (Omasits *et al.* 2014) with the domains and transmembrane topology based on alignment with the template, as well as additional secondary structure from the 3D model of the sperm whale constructed using DPSS (Kabsch & Sander 1983).

Results

Mitogenome dataset assembly

We generated mitogenome sequences for 158 new samples for this study. After combining these with 17 previously published sequences from Alexander et al. (2013), our dataset consisted of 175 globally distributed mitogenomes (Figure 1). We replicated three sequences from Alexander et al. (2013) using our capture enrichment methods and verified consistency between our approach and the long-range PCR approach of Alexander et al. (2013). Mean depth of coverage was 126× (range 18-170), and all but 13 of the newly generated sequences contained ≤10 unresolved or missing nucleotides (maximum = 65). After alignment and verification of unique indels and variant sites, there were 80 unique haplotype sequences (Table S5).

Genetic diversity analyses and relationship of haplotypes

Mitogenome nucleotide diversity ($\pi = 0.093\%$) was very similar to that reported previously for samples predominantly obtained near New Zealand (0.096%, Alexander *et al.* 2013) despite a 10-fold increase in sample size and global sample distribution. This estimate may still be slightly inflated due to selection of some specimens to maximize

371 coverage of control region haplotype diversity both in this study and in Alexander et al.
372 (2013). Sperm whales have some of the lowest documented mitogenomic diversity
373 among cetaceans (see Table 1 in Alexander *et al.* 2013), but haplotype diversity was high
374 overall (0.975) and differed significantly by ocean basin in all comparisons except "main
375 basin" Atlantic vs. Gulf of Mexico (GoMx: both haplotype and nucleotide diversity) and
376 Pacific vs. GoMx (nucleotide diversity) (Table 1; supplemental Table S6). Differences in
377 diversity remained significant for all comparisons except GoMx vs. Mediterranean
378 (haplotype diversity) even after removal of replicate haplotypes collected from the same
379 social group to control for oversampling of close relatives (supplemental Table S6).
380
381 *Phylogenetic analyses*
382 The time calibrated phylogeny for TMRCA of the four divergent sperm whale haplotypes
383 with outgroups (Phase I) is in supplemental Figure S5. The mean substitution rate
384 estimate for all codons was 7.596E-3 substitutions/site/Myr (95% CI = 6.357E-3 –
385 8.860E-3), lower than that estimated previously for the sperm whale (1.04E-2) and
386 odontocetes (1.0E-2) based on different parameters (Alexander *et al.* 2013). The median
387 TMRCA estimate of 136.7 KYA (95% CI = 85.1 – 201.1 KYA) was used as the prior for
388 analysis of the full set of sperm whale haplotypes (Phase II). The input xml, log and
389 output trees files are available in the Dryad digital archive (doi:10.5061/dryad.57271).
390 The sperm whale phylogeny based on the full set of complete mitogenome haplotypes is
391 shown in Figure 2. All equivalent sample size (ESS) values were >200 and RWTY
392 indicated convergence of topology among separate runs. The median TMRCA estimates
393 from the full-length mitogenomes and the 3rd positions of concatenated coding loci were
394 nearly identical: 126.4 KYA (95% CI = 81.04 – 180.8 KYA) and 126.5 KYA (80.066 -
395 178.4), respectively. These estimates were older than the TMRCA estimated based on 17
396 mitogenomes from the Pacific alone (103 KYA (95% HPD 72.8 - 137.4); Alexander *et*
397 *al.* 2013), likely due to the lower median substitution rate inferred in the current analysis
398 (7.034E-3 substitutions per site per million years: 95% CI = 4.068E-3 – 1.078E-2). The
399 input xml, log and output trees files are available in the Dryad digital archive
400 (doi:10.5061/dryad.57271).
401

Full length mitogenome haplotypes tended to be separated by only a few nucleotide substitutions (Figure 3). Despite this low diversity, a high degree of phylogeographic structure was evident: 65 of the 80 haplotypes were found only in the Pacific, 14 were found only in the Atlantic, and only one haplotype (mt03) was found in both of these ocean basins (1x N. Pacific, 9x N. Atlantic). Sampling in the Indian Ocean was limited to one sample from the Maldives that had a unique haplotype (mt33) that differed by 1 bp from a haplotype found in Tasmania (mt04). One additional Indian Ocean mitogenome (from the Seychelles) assembled from the genome data of Warren et al. (2017b) was haplotyped as mt54, also found in the Tasmania. The three most common control region (CR) haplotypes globally (A, B, C based on 394bp; supplemental Table S7; Alexander *et al.* 2016), which are shared between the Pacific and Atlantic Oceans (and the Mediterranean Sea), constituted 67% of the samples used in this study. These three CR haplotypes were further divided into 17, 16, and 14 mitogenome haplotypes, respectively, and the four samples from the Mediterranean Sea, all CR haplotype C, were split into two unique haplotypes (supplemental Table S7). Apart from mt03 (CR haplotype A), mentioned above, all of the 48 mitogenome haplotypes corresponding to these three CR haplotypes were ocean-basin specific.

We conducted random forest analysis to determine the probability of assigning known and newly discovered mitogenome haplotypes to ocean basin. When the GoMx and Mediterranean (Med.) samples were collapsed into the Atlantic population for purposes of assignment, results indicated 100% probability of correct assignment to the Pacific, but only 78% probability for the Atlantic Ocean (Table 2). Without the GoMx and Med. samples, the assignment probability was slightly lower for the Atlantic (71%). The large number of ocean-specific haplotypes (all but one shared Pacific/Atlantic haplotype) could suggest high diagnosability, but the lower success is due to both the single shared haplotype (mt03) and high similarity of several Atlantic haplotypes to Pacific haplotypes (Figures 2, 3).

Demographic reconstruction

The two skygrid chains showed convergence, and all skygrid parameters had ESS values above 200. The estimated TMRCA was 87 KYA (95% HPD: 60.3-119.4). The skyline analysis also showed convergence between both chains, and all parameters had ESS values >200 after combining the chains. A TMRCA very similar to the skygrid analysis was recovered (Median: 86 KYA; 95% HPD: 59-118.8 KYA). The optimal model based on AICM was the skyline. Both analyses showed patterns consistent with a population expansion (Figure 4), but the timing appeared to differ, with the skyline suggesting the expansion began 30-35 KYA and the skygrid analysis suggesting a more gradual increase from >80 KYA (Figure 4). The 95% confidence intervals for the two methods indicate little resolution in the timing of the population expansion. Since these analyses used all samples from both ocean basin populations, assumptions of the models were violated. However, when limiting the sample set to only the Atlantic or Pacific Oceans (supplemental Figure S6), the pattern of population expansion is still recovered for the Pacific, the location of the inferred mitogenome MRCA (see ancestral reconstruction section). In contrast, the skyline plot is flat for the Atlantic, as might be expected where mitogenome clusters do not coalesce within the ocean basin. Input xml and output log and tree files are available in the Dryad digital archive (doi:10.5061/dryad.57271).

Although the nuclear genome PSMC analysis also recovered evidence of a recent population expansion, it indicated that this expansion was just one in a series of population fluctuations through time. Plots from analysis of all three ocean-basin samples declined starting about 2-3 million years ago (MYA) to a low around 1 MYA, with some fluctuation through the early and mid-Pleistocene climate cycles (Figure 4). In the late Pleistocene, the low effective population size (*Ne*) approximately 120 -150 KYA, appears to have increased to a peak *Ne* roughly 50 KYA, then declined again to a low at about 20-30 KYA just before the last glacial maximum (Figure 4). This pattern is nearly identical between samples from the Atlantic, Pacific and Indian Oceans, differing only during the last glacial cycle, where the estimated population size estimate in the Atlantic was lower (though variation is higher, so differences are less certain; Supplemental Figure S1). These results are concordant with previous PSMC analyses of sperm whales (Warren *et al.* 2017b). However, the y-axis of a PSMC plot is most accurately interpreted as the

463 inverse of the coalescent rate, and changes in population structure can wrongly be
464 interpreted as changes in effective population size using this method (Foote *et al.* 2016;
465 Mazet *et al.* 2015; Mazet *et al.* 2016). To further tease apart changes in structure from
466 changes in effective population size, we generated pseudo-hybrids of pairs of individuals
467 from different populations (Cahill *et al.* 2016). As populations undergo a gradual fission,
468 the rate coalescence between the two haploid genomes that comprise the pseudo diploid
469 will decline forward in time, and should cease altogether when the two populations are
470 completely isolated (Cahill *et al.* 2016). PSMC will thus infer a gradual increase in *Ne* as
471 the coalescence rate decreases, and an infinite increase in effective population size at the
472 point in time when the lineages become completely isolated.
473
474 PSMC analyses of the pseudo-diploids indicate high uncertainty at >1 MYA, but closely
475 resemble the shared *Ne* of the three ocean-specific samples 0.15-1 MYA, indicative of a
476 stable rate of coalescence between lineages (although possibly still structured, see Chikhi
477 *et al.* 2018) during this period (Figure 4). During the Eemian inter-glacial, at around 120
478 KYA, the pseudo-diploid estimates of *Ne* increase rapidly and abruptly rises to infinity
479 for all three population-pair pseudo-diploids. We interpret this as indicative of divergence
480 between sperm whale populations associated with colonization of ocean-basins at this
481 time. Immediately following this period, estimates of *Ne* fall in both the Atlantic and
482 Pacific (and to a lesser extent, the Indian Ocean) individual PSMC plots. We interpret
483 these combined inferences as being indicative that the population split started in the
484 Eemian, likely with some ongoing gene flow (resulting in additional coalescence events
485 leading to the PSMC inferring a larger *Ne* in all samples), and that gene flow (and hence
486 coalescence) then ceased altogether after the Eemian, leading to the inferred decline in *Ne*
487 by PSMC in the individual samples. Thus, the increase and then decline in inferred *Ne* for
488 each ocean between ~125,000 and 20,000 years ago is likely to be an artifact of
489 population differentiation, masking the true effective population sizes.
490
491 *Ancestral range reconstruction and suitable habitat models*
492 Biogeographic ancestral range reconstruction analyses with BioGeoBEARS found
493 dispersal-extinction-cladeogenesis (DEC) as the most likely model, with the nested form

of the model allowing for founder-event cladogenesis (DEC+J) significantly more likely than the traditional DEC model ($p = 0.0057$). The resulting phylogeographic inference for the pruned tree indicates that the root of the tree was in the Pacific, with multiple colonizations of the Atlantic around approximately 20,000 and 60,000 years ago (Figure 5). The majority of our samples originated from the Pacific, which could potentially bias our ancestral area reconstruction (Moyle *et al.* 2016). We examined the effect of this discrepancy in sample size by down-sampling the Pacific to equal that of the Atlantic (see supplemental Figure S2). Across $n = 1{,}000$ replicates down-sampling the Pacific to 35 individuals, the root of the tree was inferred as Pacific in 99.9% of the replicates, indicating that uneven sampling is not driving this pattern.

Suitable habitat models for females in particular show a striking change between the present-day and the LGM, especially in the Atlantic Ocean (Figure 6), where core suitable habitat for females was reduced by 50% at the LGM (supplemental Table S8). The latitudinal shift also indicates that the Atlantic and Pacific habitats were likely completely separated by land masses, with only marginal potential for female dispersal between ocean basins even in the current warmer period. Projections of future habitat at the beginning of the next century suggest that dispersal between ocean basins may be more likely, but that northern and southern hemisphere populations may become separated (supplemental Figure S3).

Selection analyses on mitogenome data

PAML detected a pervasive pattern of purifying selection across the cetacean mitogenomes (supplemental Table S9), with the ratio of non-synonymous to synonymous changes estimated as $\omega = 0.093$. However, all methods except FUBAR detected at least one site putatively under positive selection in the lineage leading to sperm whales (after restricting sites inferred to be under positive selection to those where all sperm whales showed a fixed amino acid substitution in comparison with the rest of the cetacean species, supplemental Table S10). In total, 19 amino acids spanning 8 mitochondrial-encoded proteins were found to show signatures of positive selection in the sperm whale (supplemental Table S10, Figure S4). Of these 19 amino acid changes, 12 of them

525 occurred in or adjacent to transmembrane regions, contrasting with the overall pattern of
526 mammalian evolution of most adaptive variation being restricted to loop regions (da
527 Fonseca *et al.* 2008).
528
529 **Discussion**
530 Despite the high abundance and global extent of the sperm whale's distribution,
531 mitogenomic diversity within this species is markedly low (this study; Alexander *et al.*
532 2013; Lyrholm & Gyllensten 1998; Whitehead 2005). However, without detailed
533 examination of demographic patterns indicated by the nuclear genome, previous research
534 was unable to distinguish between demographic causes and selective sweeps as the most
535 likely cause of this low diversity. Here, we showed that both the mitogenome and nuclear
536 genome analyses provide evidence of a population expansion and ocean-basin divergence
537 since the last interglacial period. The presence of this pattern across both genomes is
538 consistent with a historically small effective population size (suggested by the PSMC
539 plots for all three ocean basins) rather than selective sweeps on the mitogenome or
540 cultural hitchhiking of haplotypes as the primary cause of low mtDNA diversity in sperm
541 whales, despite our detection of sites under positive selection in the sperm whale
542 mitogenome. In contrast, the expectation proposed under the mitochondrial selective
543 sweep hypotheses (due to either direct selection on the mitogenome or a matrilineal
544 cultural trait, and hitchhiking of linked neutral mtDNA diversity) is discordant patterns
545 between the nuclear and mitochondrial genomes, with a decline in mitochondrial
546 diversity over time (Whitehead 2005) but limited or no concomitant decline in nuclear
547 DNA diversity. This expectation is at least partially due to the high rates of sex-specific
548 dispersal of males in sperm whales (Alexander et al. 2016), and recombination in the
549 nuclear genome, unlinking the nuclear and mitochondrial genomes in these processes.
550 However, it is important to note that given the strong influence of social structure on
551 sperm whales (Whitehead *et al.* 2017), and the detection of positive selection on the
552 protein-coding regions, we cannot rule out other forces, including cultural hitchhiking,
553 having further reduced mitogenome diversity in the sperm whale. Nevertheless, the
554 consistency of the inferred reduction in population size based on the nuclear genome
555 occurring at approximately the same time as the TMRCA of the mitogenomes suggests

556 demographic processes as the primary cause of low present-day sperm whale mtDNA
557 diversity.
558
559 Our mitogenomic analyses suggest that the current global distribution of sperm whales
560 results from a relatively recent expansion (20-40 KYA). The inferred mitogenomic
561 MRCA suggests expansion from a single, refugial population most likely located in the
562 Pacific Ocean (though we cannot rule out an Indian Ocean refugium due to lack of
563 sampling). "Ice-house" conditions are believed to have developed approximately 3 MYA,
564 characterized by the development of perennial Arctic Ocean sea ice and continental ice
565 sheets in North America and Eurasia (Greene *et al.* 2008). This global change has been
566 implicated in a global extinction event among marine megafauna (Pimiento *et al.* 2017)
567 and coincides with an inferred decline in the global sperm whale population (revealed by
568 PSMC analyses of individual sperm whale genomes), and an evident restriction in the
569 distribution of sperm whales (revealed by pseudo-diploid PSMC analysis of the nuclear
570 genome, Figure 4), potentially reducing the distribution to a single refugial population
571 from which present day sperm whale lineages descended. The global abundance appears
572 to have increased from a low of <10,000 breeding individuals (*Ne*) since the last long
573 cold period (Saale glaciation, ~80 kyr) that ended approximately 125 KYA. Since then,
574 there have been several shorter cold periods, lasting <10 kyr each (Figure 4), which
575 appear to correspond approximately with coalescence nodes in the mitogenome tree
576 (Figure 2), after which haplotypes diversified and mitochondrial lineages dispersed
577 between oceans. Although timing estimates are approximate, the pattern suggests effects
578 of climate on population size, distribution, and persistence of specific maternal lineages
579 through time in the sperm whale. While male sperm whales are known to travel to and
580 feed in cold temperate/polar waters, females are generally restricted to warmer tropical
581 and temperate waters, and mtDNA data suggest that they may only be able to disperse
582 between the Pacific and Atlantic oceans during warm periods that allow them to extend
583 their latitudinal range. However, the pseudo-diploid analysis and divergence of the
584 nuclear genome PSMC plots from samples located in different oceans over the last
585 ~100,000 years (Figure 4) both suggest that males also rarely disperse between ocean
586 basins, resulting in genetic divergence between the Atlantic and Pacific (see also

587 Alexander *et al.* 2016; Lyrholm & Gyllensten 1998; Lyrholm *et al.* 1999). Nuclear
588 genome-wide SNP analysis of more samples from each ocean basin – particularly the
589 under-sampled Indian Ocean – will be needed to infer degree of divergence and levels of
590 current and/or historical gene flow.
591
592 The use of complete mitogenomes instead of control region sequences allows for
593 substantially greater power to detect phylogeographic structure in low-diversity species
594 such as the sperm whale, and can also improve power in population genetic studies. The
595 sampling design for this study was not appropriate for traditional frequency-based
596 population genetic analyses, but comparative analysis of mitogenome and control region
597 haplotypes for large geographic regions (supplemental Figure S7) indicated mitogenomes
598 were more sensitive at detecting significant differentiation among strata based on F_{ST},
599 Φ_{ST} and χ^2 statistics. Although patterns of differentiation were largely concordant
600 between these data sets, the lack of power from control region data indicates caution must
601 be used when genetic structure among potential strata is not detected solely based on CR
602 data.
603
604 Sperm whales are among the deepest diving marine mammals, and it is likely that their
605 mitochondrially-encoded proteins have been under intense selection for oxygen
606 efficiency (Janik 2001) and robustness to pressure changes (Somero 1992; Warren *et al.*
607 2017b). Changes in transmembrane regions could more strongly affect protein structure
608 and function than in the loop regions (Saier 1994), potentially directly addressing these
609 selective pressures. Future comparisons to patterns of positive selection on other deep-
610 diving species would be needed to see if this was a pervasive pattern across other taxa
611
612 Sperm whales are not considered rare, but have been depleted to approximately 1/3 of
613 their pre-whaling population sizes and are protected by various national and international
614 laws and treaties (e.g., the US Endangered Species Act, IUCN Red List, International
615 Whaling Commission). Given evidence of population structure within and between ocean
616 basins (Alexander *et al.* 2016; Engelhaupt *et al.* 2009; Lyrholm & Gyllensten 1998;
617 Mesnick *et al.* 2011), the likelihood that populations were unevenly depleted by whaling

(Ivashchenko *et al.* 2014), and lack of information on population recovery globally, it is highly probable that some sperm whale populations are still endangered or at risk (Carroll *et al.* 2014; Gero & Whitehead 2016; Notarbartolo-Di-Sciara 2014). Our results provide support for current isolation of populations located in different ocean basins, with episodic dispersal of females restricted to only warm climate periods. Although the current warming trends predict expansion of sperm whale habitat that allows both growth and opportunities for inter-ocean dispersal, we cannot predict how rapid climate change may affect the ecosystems on which sperm whales depend (see, e.g., Dawe *et al.* 2007; Ibanez *et al.* 2011; Jaquet *et al.* 2003; Pecl & Jackson 2008). In particular, climate change can affect the Atlantic and Pacific oceans quite differently (Boyle & Keigwin 1985; Cheng *et al.* 2009; Greene *et al.* 2008; Howard 1997), which may explain the apparent refugial population of sperm whales in the Pacific Ocean during the last long glacial period. Further research on trends in abundance and health of sperm whale populations and protection from other anthropogenic effects such as pollution (e.g., de Stephanis *et al.* 2013; Savery *et al.* 2014a; Savery *et al.* 2014b; Unger *et al.* 2016), ocean noise (Mate *et al.* 1994), ship strikes (Jensen & Silber 2003), entanglement (Barlow & Cameron 2003), and prey competition (Hucke-Gaete *et al.* 2004) should be top priorities to prevent loss of already depleted populations.

Acknowledgements

We are grateful to all who collected and contributed samples used in this study: Marie Hill of the PIFSC Cetacean Research Program and Joint Institute for Marine and Atmospheric Research, Erin Oleson (PIFSC), Allan Ligon, Adam Ü, Mark Deakos, the Commander, U.S. Pacific Fleet, Office of Naval Research and Living Marine Resources Programs, Diane Claridge, John Durban, Kim Parsons, Luke Rendell, Barb Taylor, Louella Dolar, William Perrin, Deborah Thiele, Karen Evans, Robert Pitman, Tim Gerrodette, Tom Jefferson, Samual Hung, Liz Slooten, Sophie Laran, Emilie Praca, Lisa Ballance, Paul Wade, William McLellan, Ken Balcomb, Aleta Hohn, Jay Barlow, Dan Odell, Hisao Azuma, Christine Adkins, Wayne Haggard, Lars Kleivane, Manuel Furtado Neto, Sue Barco, Wayne McFee, K. Mullin, and the Southeast US marine mammal

stranding network. All samples have been imported to the US under CITES permit, and samples from the US were collected under Marine Mammal Protection Act permits. We are grateful for technical and analytical assistance from The Scripps Research Institute core sequencing lab, Sebastian Duchene, Simon Ho, D. Steel, K. Hoekzema, C. Sislak and Milaya Nykänen. Cristina Garilao generously provided customized distribution maps from AquaMaps. We are very grateful for helpful discussion and comments on the manuscript provided by Hal Whitehead, Kim Parsons, Karen Chambers and four anonymous reviewers. Funding was provided by the National Marine Fisheries Service Southwest Region; the Marine Mammal and Turtle Division, Southwest Fisheries Science Center; U.S. Navy Pacific Fleet and NMFS Pacific Islands Science Center. Financial support for ADF was provided by the Welsh Government and Higher Education Funding Council for Wales through the Sêr Cymru National Research Network for Low Carbon, Energy and Environment, and from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 663830.

References

Alexander A, Steel D, Hoekzema K, Mesnick SL, Engelhaupt D, Kerr I, . . . Baker CS (2016) What influences the worldwide genetic structure of sperm whales (*Physeter macrocephalus*)? *Molecular Ecology* **25**, 2754-2772.

Alexander A, Steel D, Slikas B, Hoekzema K, Carraher C, Parks M, . . . Baker CS (2013) Low diversity in the mitogenome of sperm whales revealed by next-generation sequencing. *Genome Biology and Evolution* **5**, 113-129.

Amos W (1999) Culture and genetic evolution in whales. *Science* **284**, 2055a.

Archer FI, Adams PE, Schneiders BB (2017a) strataG: An R package for manipulating, summarizing, and analyzing population genetic data. *Molecular Ecology Resources* **17**, 5-11.

Archer FI, Martien KK, Taylor BL (2017b) Diagnosability of mtDNA with Random Forests: Using sequence data to delimit subspecies. *Marine Mammal Science* **33**, 101-131.

Archer FI, Morin PA, Hancock-Hanser BL, Robertson KM, Leslie MS, Berube M, . . . Taylor BL (2013) Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus* spp.): genetic evidence for revision of subspecies. *PLoS ONE* **8**, e63396.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison

while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* **29**, 2157-2167.

Baker CS, Slade RW, Bannister JL, Abernethy RB, Weinrich MT, Lien J, . . . Palumbi SR (1994) Hierarchical structure of mitochondrial DNA gene flow among humpback whales Megaptera novaeangliae, world-wide. *Molecular Ecology* **3**, 313-327.

Baker CS, Steel D, Calambokidis J, Falcone E, Gonzalez-Peral U, Barlow J, . . . Yamaguchi M (2013) Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology Progress Series* **494**, 291-306.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37-48.

Barlow J, Cameron GA (2003) Field experiments show that acoustic pingers reduce marine mammal bycatch in the California drift gill net fishery. *Marine Mammal Science* **19**, 265-283.

Bazin E, Glemin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**, 570-572.

Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, . . . Drummond AJ (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* **10**, e1003537.

Boyle EA, Keigwin LD (1985) Comparison of Atlantic and Pacific paleochemical records for the last 215,000 years - changes in deep ocean circulation and chemical inventories. *Earth and Planetary Science Letters* **76**, 135-150.

Breiman L (2001) Random forests. *Machine Learning* **45**, 5-32.

Buddhakosai W, Klinsawat W, Smith O, Sukmak M, Kaolim N, Duangchantrasiri S, . . . Wajjwalku W (2016) Mitogenome analysis reveals a complex phylogeographic relationship within the wild tiger population of Thailand. *Endangered Species Research* **30**, 125-131.

Caballero S, Duchene S, Garavito MF, Slikas B, Baker CS (2015) Initial evidence for adaptive selection on the NADH subunit two of freshwater dolphins by analyses of mitochondrial genomes. *PLoS ONE* **10**, e0123543.

Cahill JA, Soares AE, Green RE, Shapiro B (2016) Inferring species divergence times using pairwise sequential Markovian coalescent modelling and low-coverage genomic data. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* **371**.

Carroll G, Hedley S, Bannister J, Ensor P, Harcourt R (2014) No evidence for recovery in the population of sperm whale bulls off Western Australia, 30 years post-whaling. *Endangered Species Research* **24**, 33-43.

Cheng H, Edwards RL, Broecker WS, Denton GH, Kong XG, Wang YJ, . . . Wang XF (2009) Ice age terminations. *Science* **326**, 248-252.

Chikhi L, Rodriguez W, Grusea S, Santos P, Boitard S, Mazet O (2018) The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. *Heredity (Edinb)* **120**, 13-24.

da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A (2008) The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics* **9**, 119.

Darriba D, Taboada G, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**, 772.

Dawe EG, Hendrickson LC, Colbourne EB, Drinkwater KF, Showell MA (2007) Ocean climate effects on the relative abundance of short-finned (*Illex illecebrosus*) and long-finned (*Loligo pealeii*) squid in the northwest Atlantic Ocean. *Fisheries Oceanography* **16**, 303-316.

de Stephanis R, Gimenez J, Carpinelli E, Gutierrez-Exposito C, Canadas A (2013) As main meal for sperm whales: plastics debris. *Marine Pollution Bulletin* **69**, 206-214.

Delport W, Poon AF, Frost SD, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26**, 2455-2457.

Dornburg A, Brandley MC, McGowen MR, Near TJ (2012) Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea). *Molecular Biology and Evolution* **29**, 721-736.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology & Evolution* **29**, 1969-1973.

Engelhaupt D, Hoelzel AR, Nicholson C, Frantzis A, Mesnick S, Gero S, . . . Mignucci-Giannoni AA (2009) Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (*Physeter macrocephalus*). *Molecular Ecology* **18**, 4193-4205.

Finch TM, Zhao N, Korkin D, Frederick KH, Eggert LS (2014) Evidence of positive selection in mitochondrial complexes I and V of the African elephant. *PLoS ONE* **9**, e92587.

Foote AD, Morin PA, Durban JW, Pitman RL, Wade P, Willerslev E, . . . da Fonseca RR (2011) Positive selection on the killer whale mitogenome. *Biology Letters* **7**, 116-118.

Foote AD, Vijay N, Avila-Arcos MC, Baird RW, Durban JW, Fumagalli M, . . . Wolf JB (2016) Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *Nature Communications* **7**, 11693.

Gero S, Whitehead H (2016) Critical decline of the eastern caribbean sperm whale population. *PLoS ONE* **11**.

Greene CH, Pershing AJ, Cronin TM, Ceci N (2008) Arctic climate change and its impacts on the ecology of the North Atlantic. *Ecology* **89**, S24-38.

Guindon S, Delsuc F, Dufayard JF, Gascuel O (2009) Estimating maximum likelihood phylogenies with PhyML. *Methods in Molecular Biology* **537**, 113-137.

Hancock-Hanser B, Frey A, Leslie M, Dutton PH, Archer EI, Morin PA (2013) Targeted multiplex next-generation sequencing: Advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics *Molecular Ecology Resources* **13**, 254-268.

Heller R, Chikhi L, Siegismund HR (2013) The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS ONE* **8**.

Ho SY, Lanfear R (2010) Improved characterisation of among-lineage rate variation in cetacean mitogenomes using codon-partitioned relaxed clocks. *Mitochondrial DNA* **21**, 138-146.

Hoelzel AR, Natoli A, Dahlheim ME, Olavarria C, Baird RW, Black NA (2002) Low worldwide genetic diversity in the killer whale (*Orcinus orca*): implications for demographic history. *Proceedings of the Royal Society B-Biological Sciences* **269**, 1467-1473.

Howard WR (1997) Palaeoclimatology - A warm future in the past. *Nature* **388**, 418-419.

Hucke-Gaete R, Moreno CA, Arata J, Ctr BW (2004) Operational interactions of sperm whales and killer whales with the Patagonian toothfish industrial fishery off southern Chile. *Ccamlr Science* **11**, 127-140.

Ibanez CM, Cubillos LA, Tafur R, Arguelles J, Yamashiro C, Poulin E (2011) Genetic diversity and demographic history of *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the Humboldt Current System. *Marine Ecology Progress Series* **431**, 163-171.

Ivashchenko YV, Brownell RL, Clapham PJ (2014) Distribution of Soviet catches of sperm whales *Physeter macrocephalus* in the North Pacific. *Endangered Species Research* **25**, 249-263.

Jackson JA, Steel DJ, Beerli P, Congdon BC, Olavarria C, Leslie MS, . . . Baker CS (2014) Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). *Proceedings of the Royal Society B-Biological Sciences* **281**.

Janik VM (2001) Is cetacean social learning unique? *Behavioral and Brain Sciences* **24**, 337-338.

Jaquet N, Gendron D, Coakes A (2003) Sperm whales in the Gulf of California: Residency, movements, behavior, and the possible influence of variation in food supply. *Marine Mammal Science* **19**, 545-562.

Jaquet N, Whitehead H (1996) Scale-dependent correlation of sperm whale distribution with environmental features and productivity in the South Pacific. *Marine Ecology Progress Series* **135**, 1-9.

Jensen AS, Silber GK (2003) Large Whale Ship Strike Database, p. 37. National Marine Fisheries Service, NOAA Technical Memorandum NMFS-OPR-25, Silver Spring, MD.

Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637.

Kaschner K, Kesner-Reyes K, Garilao C, Rius-Barile J, Rees T, Froese R (2016) AquaMaps: Predicted range maps for aquatic species. *World wide web electronic publication, www.aquamaps.org, Version 08/2016.*

Kaschner K, Tittensor DP, Ready J, Gerrodette T, Worm B (2011) Current and future patterns of global marine mammal biodiversity. *PLoS ONE* **6**, e19653.

Kopps AM, Ackermann CY, Sherwin WB, Allen SJ, Bejder L, Krützen M (2014) Cultural transmission of tool use combined with habitat specializations leads to fine-scale genetic structure in bottlenose dolphins. *Proc Biol Sci* **281**, 20133245.

Lanfear R, Calcott B, Ho S, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29**, 1695-1710.

Lang AR, Calambokidis J, Scordino J, Pease VL, Klimek A, Burkanov VN, . . . Taylor BL (2014) Assessment of genetic structure among eastern North Pacific gray whales on their feeding grounds. *Marine Mammal Science* **30**, 1473-1493.

Leduc RG, Archer FI, Lang AR, Martien KK, Hancock-Hanser B, Torres-Florez JP, . . . Taylor BL (2016) Genetic variation in blue whales in the eastern pacific: implication for taxonomy and use of common wintering grounds. *Molecular Ecology*.

Leigh JW, Bryant D (2015) POPART: fullfeature software for haplotype network construction. *Methods in Ecology and Evolution* **6**, 1110-1116.

Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493-496.

Lyrholm T, Gyllensten U (1998) Global matrilineal population structure in sperm whales as indicated by mitochondrial DNA sequences. *Proceedings of the Royal Society B-Biological Sciences* **265**, 1679-1684.

Lyrholm T, Leimar O, Gyllensten U (1996) Low diversity and biased substitution patterns in the mitochondrial DNA control region of sperm whales: implications for estimates of time since common ancestry. *Molecular Biology and Evolution* **13**, 1318-1326.

Lyrholm T, Leimar O, Johanneson B, Gyllensten U (1999) Sex-biased dispersal in sperm whales: contrasting mitochondrial and nuclear genetic structure of global populations. *Proceedings of the Royal Society B-Biological Sciences* **266**, 347-354.

Martien KK, Chivers SJ, Baird RW, Archer FI, Gorgone AM, Hancock-Hanser BL, . . . Taylor BL (2014) Nuclear and mitochondrial patterns of population structure in north pacific false killer whales (*Pseudorca crassidens*). *Journal of Heredity* **105**, 611-626.

Mate BR, Stafford KM, Ljungblad DK (1994) A change in sperm whale (*Physeter macrocephalus*) distribution correlated to seismic surveys in the Gulf of Mexico. *Journal of the Acoustical Society of America* **96**, 3268-3269.

Matzke NJ (2012) Founder-event speciation in BioGeoBEARS package dramatically improves likelihoods and alters parameter inference in Dispersal-Extinction-Cladogenesis (DEC) analyses. *Frontiers of Biogeography* **4(suppl. 1)**, 210.

Matzke NJ (2013) *BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts*, University of California, Berkeley.

Matzke NJ (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology* **63**, 951-970.

Mazet O, Rodriguez W, Chikhi L (2015) Demographic inference using genetic data from a single individual: Separating population size variation from population structure. *Theoretical Population Biology* **104**, 46-58.

Mazet O, Rodriguez W, Grusea S, Boitard S, Chikhi L (2016) On the importance of being structured: instantaneous coalescence rates and human evolution--lessons for ancestral population size inference? *Heredity (Edinb)* **116**, 362-371.

McGowen MR, Spaulding M, Gatesy J (2009) Divergence date estimation and a comprehensive molecular tree of extant cetaceans. *Molecular Phylogenetics and Evolution* **53**, 891-906.

Mesnick S, Taylor B, Archer EI, Martien K, Escorza Treviño S, Hancock-Hanser BL, . . . Morin PA (2011) Sperm whale population structure in the eastern North Pacific inferred by the use of single nucleotide polymorphisms (SNPs), microsatellites and mitochondrial DNA. *Molecular Ecology Resources* **11 (suppl. 1)**, 278-298.

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16**, 1215.

Mizroch SA, Rice DW (2013) Ocean nomads: Distribution and movements of sperm whales in the North Pacific shown by whaling data and Discovery marks. *Marine Mammal Science* **29**, E136-E165.

Morin PA, Archer FI, Foote AD, Vilstrup J, Allen EE, Wade P, . . . Harkins T (2010) Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Research* **20**, 908-916.

Morin PA, Parsons KM, Archer FI, Ávila-Arcos MC, Barrett-Lennard LG, Dalla Rosa L, . . . Foote AD (2015) Geographic and temporal dynamics of a global radiation and diversification in the killer whale. *Molecular Ecology* **24**, 3964-3979.

Moyle RG, Oliveros CH, Andersen MJ, Hosner PA, Benz BW, Manthey JD, . . . Faircloth BC (2016) Tectonic collision and uplift of Wallacea triggered the global songbird radiation. *Nature Communications* **7**, 12709.

Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolution* **30**, 1196-1205.

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics* **8**, e1002764.

Notarbartolo-Di-Sciara G (2014) Sperm whales, *Physeter macrocephalus*, in the Mediterranean Sea: a summary of status, threats, and conservation recommendations. *Aquatic Conservation-Marine and Freshwater Ecosystems* **24**, 4-10.

Nylander JA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**, 581-583.

O'Brien SJ (1994) A role for molecular genetics in biological conservation. *Proceedings of the National Academy of Science USA* **91**, 5748-5755.

Omasits U, Ahrens CH, Muller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30**, 884-886.

Pecl GT, Jackson GD (2008) The potential impacts of climate change on inshore squid: biology, ecology and fisheries. *Reviews in Fish Biology and Fisheries* **18**, 373-385.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605-1612.

Pimiento C, Griffin JN, Clements CF, Silvestro D, Varela S, Uhen MD, Jaramillo C (2017) The Pliocene marine megafauna extinction and its impact on functional diversity. *Nat Ecol Evol* **1**, 1100-1106.

912 Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on
913 individual sites of codon alignments. *Bioinformatics* **21**, 2531-2533.
914 Pond SL, Frost SD, Muse SV (2005) HyPhy: hypothesis testing using phylogenies.
915 *Bioinformatics* **21**, 676-679.
916 Premo LS, Hublin JJ (2009) Culture, population structure, and low genetic diversity in
917 Pleistocene hominins. *Proc Natl Acad Sci U S A* **106**, 33-37.
918 R Core Team (2014) R: A language and environment for statistical computing. R
919 Foundation for Statistical Computing, Vienna, Austria. https://www.R-
920 project.org.
921 Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6.
922 http://tree.bio.ed.ac.uk/software/tracer/.
923 Ready J, Kaschner K, South AB, Eastwood PD, Rees T, Rius J, . . . Froese R (2010)
924 Predicting the distributions of marine organisms at the global scale. *Ecological*
925 *Modelling* **221**, 467-478.
926 Saier MH (1994) Computer-aided analyses of transport protein sequences: gleaning
927 evidence concerning function, structure, biogenesis, and evolution.
928 *Microbiological Reviews* **58**, 71-93.
929 Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*,
930 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
931 Savery LC, Wise JT, Wise SS, Falank C, Gianios C, Jr., Thompson WD, . . . Wise JP, Sr.
932 (2014a) Global assessment of arsenic pollution using sperm whales (*Physeter*
933 *macrocephalus*) as an emerging aquatic model organism. *Comparative*
934 *Biochemistry and Physiology. Toxicology and Pharmacology.* **163**, 55-63.
935 Savery LC, Wise SS, Falank C, Wise J, Gianios C, Jr., Douglas Thompson W, . . . Wise
936 JP, Sr. (2014b) Global assessment of oceanic lead pollution using sperm whales
937 (*Physeter macrocephalus*) as an indicator species. *Marine Pollution Bulletin* **79**,
938 236-244.
939 Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein
940 homology-modeling server. *Nucleic Acids Research* **31**, 3381-3385.
941 Shamblin BM, Bjorndal KA, Bolten AB, Hillis-Starr ZM, Lundgren I, Naro-Maciel E,
942 Nairn CJ (2012) Mitogenomic sequences better resolve stock structure of southern
943 Greater Caribbean green turtle rookeries. *Molecular Ecology* **21**, 2330-2340.
944 Somero GN (1992) Biochemical ecology of deep-sea animals. *Experientia* **48**, 537-543.
945 Steiner L, Lamoni L, Plata MA, Jensen SK, Lettevall E, Gordon J (2012) A link between
946 male sperm whales, *Physeter macrocephalus*, of the Azores and Norway. *Journal*
947 *of the Marine Biological Association of the United Kingdom* **92**, 1751-1756.
948 Straley JM, Schorr GS, Thode AM, Calambokidis J, Lunsford CR, Chenoweth EM, . . .
949 Andrews RD (2014) Depredating sperm whales in the Gulf of Alaska: local
950 habitat use and long distance movements across putative population boundaries.
951 *Endangered Species Research* **24**, 125-135.
952 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular
953 Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**,
954 2725-2729.
955 Taylor BL, Chivers SJ, Larese J, Perrin WF (2007) Generation length and percent mature
956 estimates for IUCN assessments of cetaceans. Administrative Report, Southwest
957 Fisheries Science Center, 8604 La Jolla Shores Blvd., La Jolla, CA 92038, USA.

Tiedemann R, Milinkovitch MC (1999) Culture and genetic evolution in whales. *Science* **284**, 2055a.

Unger B, Rebolledo EL, Deaville R, Grone A, LL IJ, Leopold MF, . . . Herr H (2016) Large amounts of marine debris found in sperm whales stranded along the North Sea coast in early 2016. *Marine Pollution Bulletin* **112**, 134-141.

Van Cise A, Morin PA, Baird RW, Lang AR, Robertson KM, Chivers SJ, . . . Martien KK (2016) Redrawing the map: mtDNA provides new insights into the distribution and diversity of short-finned pilot whales in the Pacific Ocean. *Marine Mammal Science* **32**, 1177-1199.

Warren DL, Geneva AJ, Lanfear R (2017a) RWTY (R We There Yet): An R package for examining convergence of Bayesian phylogenetic analyses. *Molecular Biology and Evolution* **34**, 1016-1020.

Warren WC, Kuderna L, Alexander A, Catchen J, Pérez-Silva JG, López-Otín C, . . . Wise Sr. JP (2017b) The novel evolution of the sperm whale genome. *Genome Biology and Evolution* **9**, 3260-3264.

Whitehead H (1998) Cultural selection and genetic diversity in matrilineal whales. *Science* **282**, 1708-1711.

Whitehead H (2002) Estimates of the current global population size and historical trajectory for sperm whales. *Marine Ecology Progress Series* **242**, 295-304.

Whitehead H (2005) Genetic diversity in the matrilineal whales: Models of cultural hitchhiking and group-specific non-heritable demographic variation. *Marine Mammal Science* **21**, 58-79.

Whitehead H, Vachon F, Frasier TR (2017) Cultural hitchhiking in the matrilineal whales. *Behavior Genetics* **47**, 324-334.

Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://king2.sc.fsu.edu/CEBProjects/awty/awty_start.php.

Woolley S, Johnson J, Smith MJ, Crandall KA, McClellan DA (2003) TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics* **19**, 671-672.

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**, 1586-1591.

Yang Z, Wong WS, Nielsen R (2005) Bayes empirical bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* **22**, 1107-1118.

Data Accessibility

Mitogenome haplotype sequences generated for this study have been submitted to GenBank (accession numbers KU891329 - KU891394). Sequence alignment input and output xml file from Beast phylogenetic and skyline/skygrid analyses are available from the Dryad Digital Repository doi:10.5061/dryad.57271.

Author Contributions

PAM and SLM conceived of the study. CSB, BRM, SLM, HW, KK, PER and AA contributed samples, data, and expert interpretation. BLH, VLP and AA generated genetic data. PAM, AA and ADF conducted analyses and wrote the manuscript.

Figures legends

Figure 1. Map of all samples used in this study. Black circles = new mitogenome data, gray triangles = previously published mitogenome sequences (Alexander *et al.* 2013).

Figure 2. Time-calibrated (millions of years) phylogeny of all unique sperm whale mitochondrial genomes. Posterior probability support >0.5 is shown at nodes. Haplotype ID's are provided as "mt##" (see supplemental Table S1 for samples associated with haplotypes) followed by abbreviated locations (see supplemental Table S1). The clades to the right correspond to the pruned tree clades used for phylogeographic analysis (Figure 5). Branches are color coded by ocean basin where haplotypes are found: Blue = Pacific, Red = Atlantic, Green = Indian, Orange = Mediterranean, Pink = Pacific/Atlantic (shared). *Mitogenome haplotypes assembled from SRA data (Warren et al. 2017b).

Figure 3. Median Joining Network of all unique haplotypes, colored by ocean. Tick marks on branches indicate the number of nucleotide differences. Not all 80 unique haplotypes are included in the network, as the algorithm collapses haplotypes with missing or ambiguous positions.

Figure 4. Demographic reconstructions based on the sperm whale mitogenomes and nuclear genome PSMC. The mitogenome skyline (black) and skygrid (gray) plot lower and upper 95% highest posterior density intervals are shown by thinner lines and median values by thicker lines. Demographic estimates were converted to effective population size (females) by dividing by a τ of 31.9 years (Taylor *et al.* 2007). The nuclear PSMC plots are for three sperm whales from the Atlantic (red), Pacific (blue), and Indian (yellow) Oceans. Pseudo-diploid plots are shown for each ocean pair: Atlantic-Pacific (green), Atlantic-Indian (pink), and Pacific-Indian (purple). The X axis starts at 10,000 yr before present, and the Y axis is truncated at $3.5x10^4$, as pseudo-diploid plots increase exponentially and bootstrap variation at $<2x10^4$ years was very large. Glacial maxima and Eemian warm period are shown with beige and yellow shading, respectively. Bootstrap plots for each PSMC analysis are presented in supplemental Figure S1.

Figure 5. Most probable ancestral ranges for sperm whale populations estimated using the best fit model DEC+J from BioGeoBEARS and full mitogenomes. Clade IDs (tips) are as in Figure 2, and represent the tree topology pruned to one tip per clade. The boxes at nodes indicate the most likely geographic range immediately after cladogenesis. Oceanic regions are indicated by P = Pacific, A = Atlantic, I = Indian, M = Mediterranean.

Figure 6. AquaMaps environmental envelope models for distribution of female sperm whale habitat mapped to environmental conditions for a) last glacial maximum and b) current habitat. Dark red color indicates core suitable habitat; yellow indicates marginal habitat.

Table 1
Number of haplotypes, haplotypic diversity and nucleotide diversity by ocean basin.

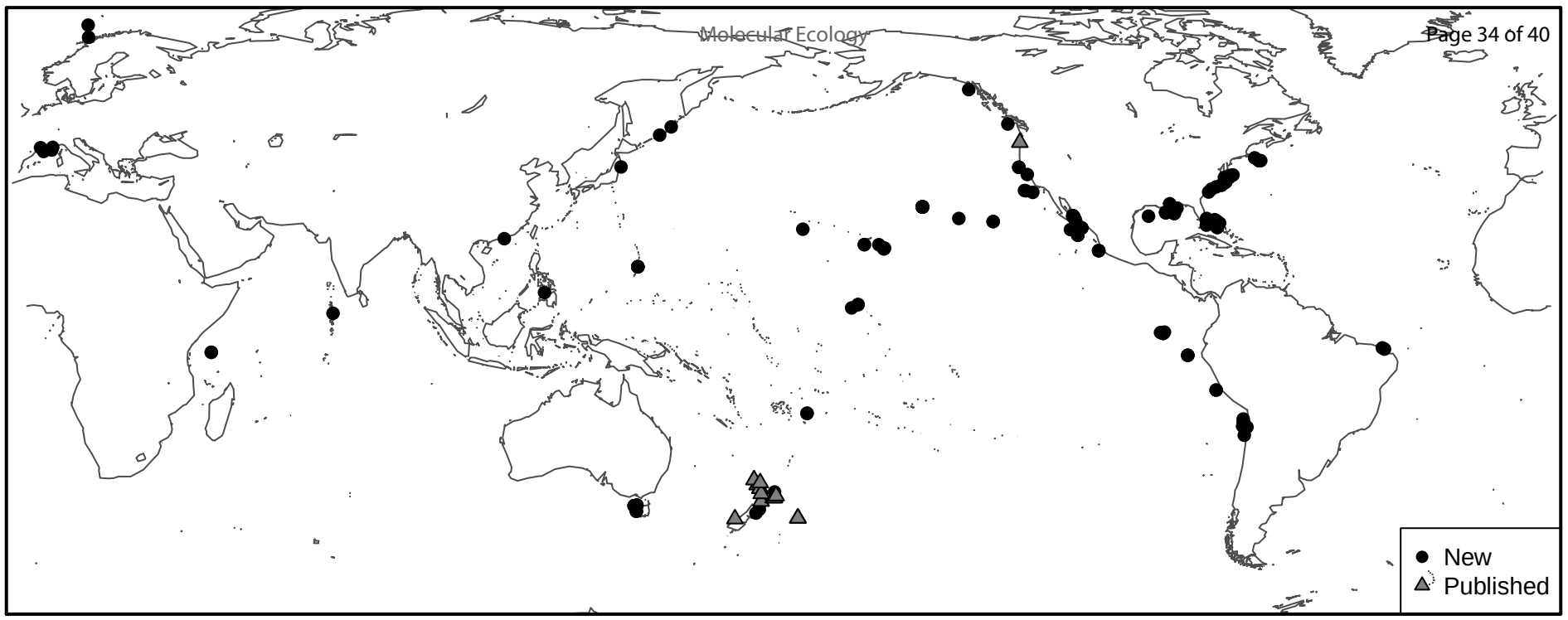
	No. Samples	No. Haplotypes	Haplotype Diversity	Nucleotide Diversity (π)
Atlantic*	35	10	0.839	0.000991
GoMx	11	5	0.818	0.001019
Mediterranean	4	2	0.500	0.000000
Pacific	124	66	0.972	0.000795
Global†	175	80	0.975	0.000934

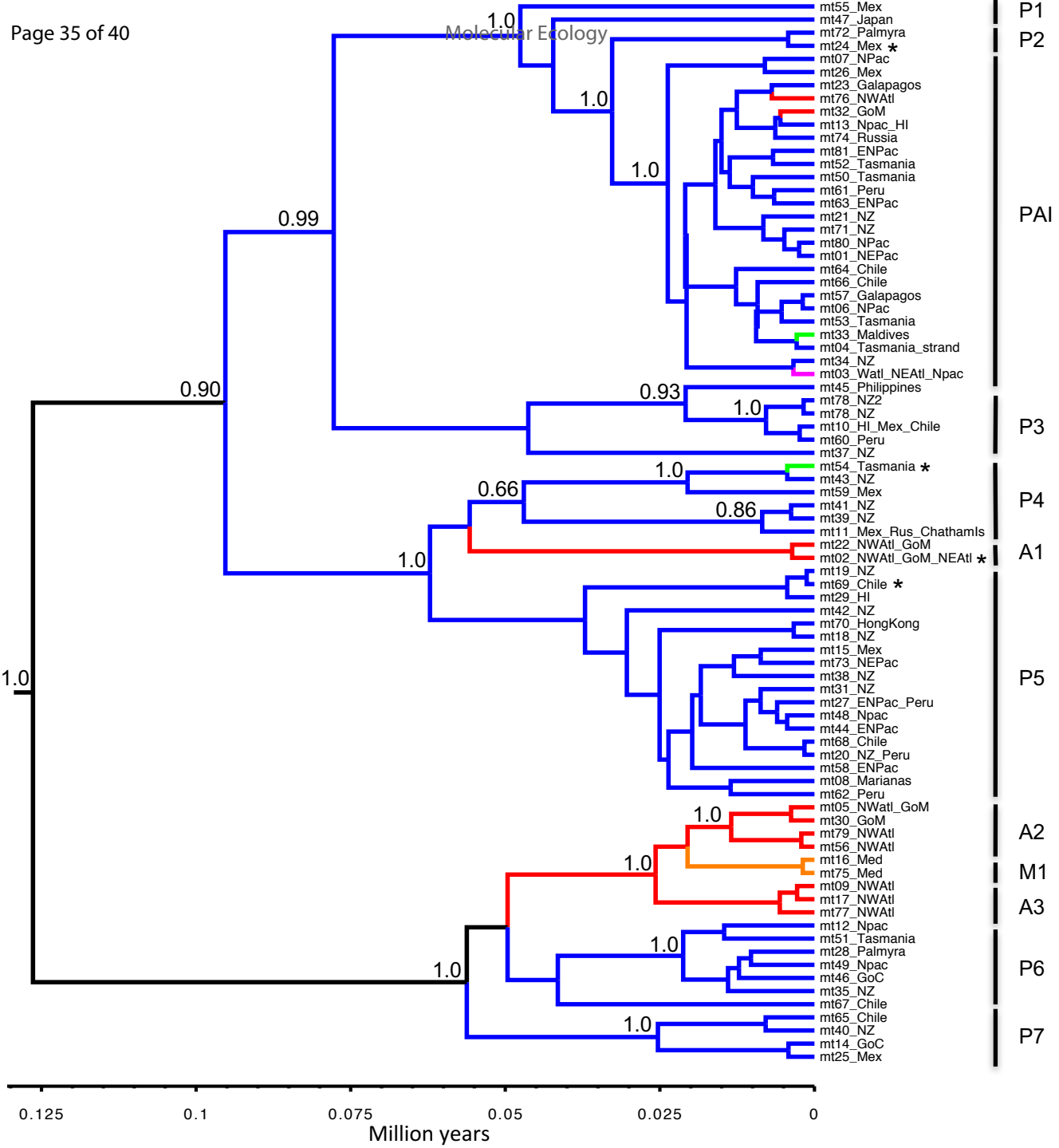
*excluding the Gulf of Mexico and Mediterranean Sea.

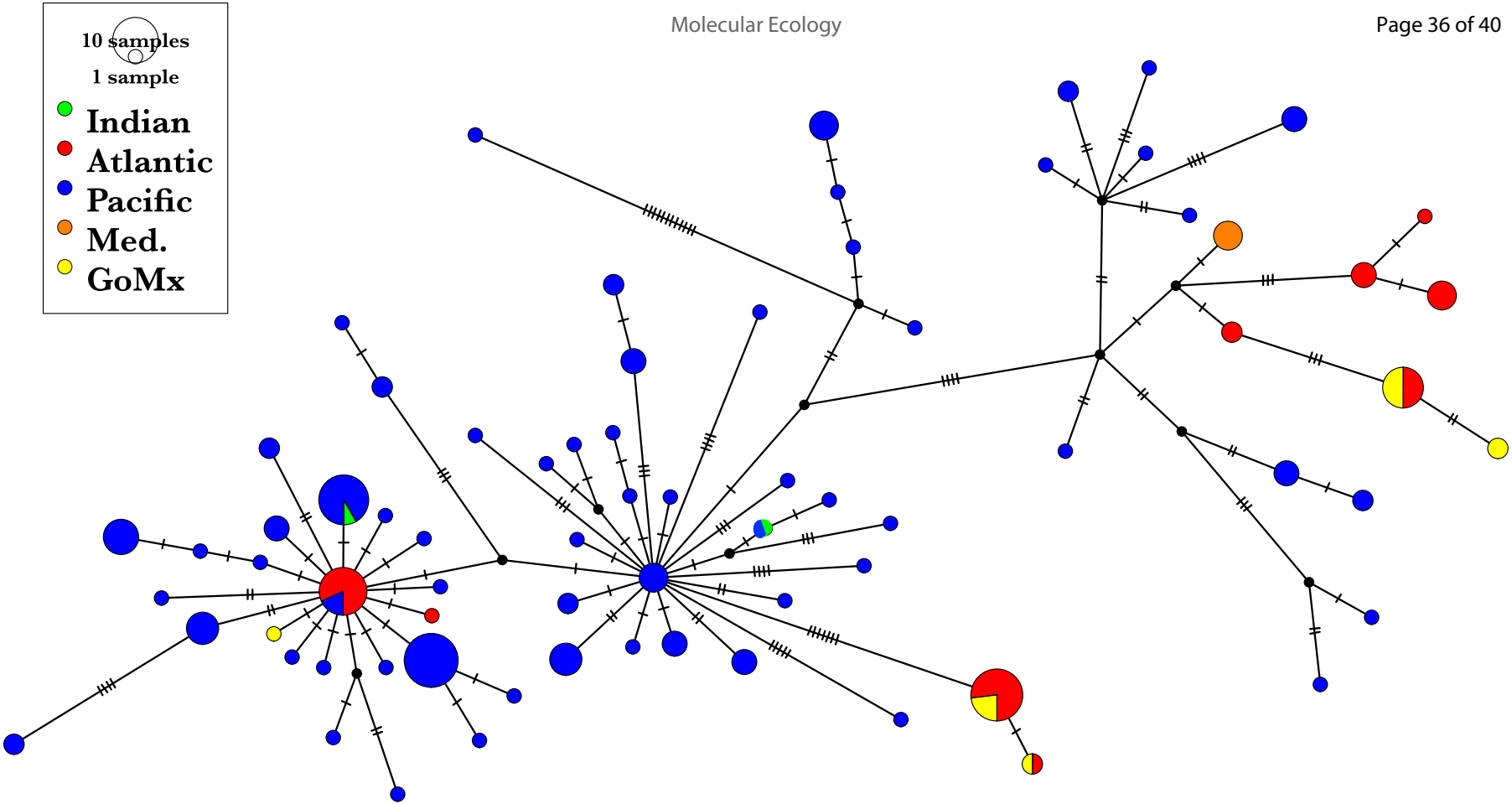
†includes single sample from the Indian Ocean.

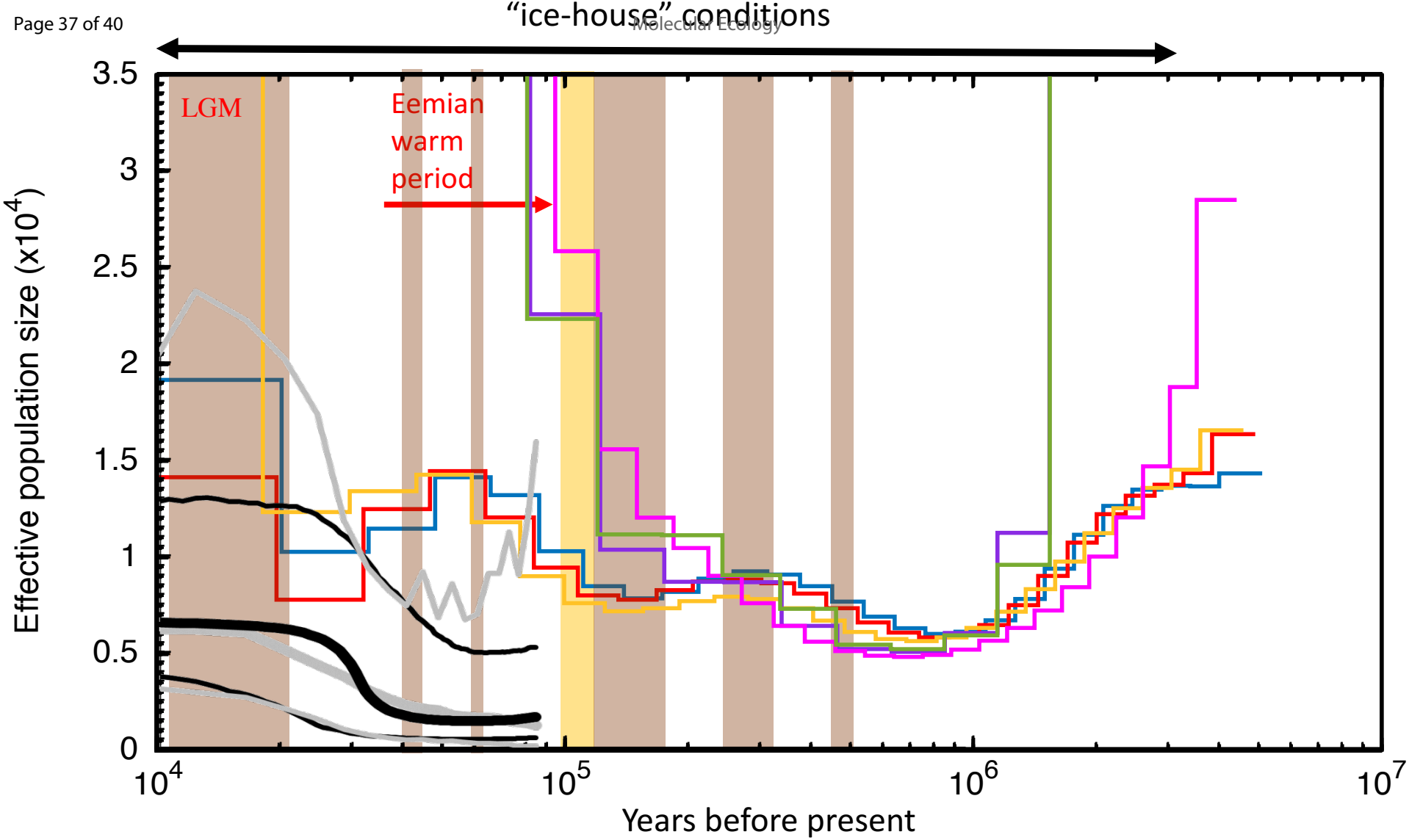
Table 2
Random forest assignment of samples to ocean basin, percent correctly assigned, and lower (LCI) and upper (UCI) 95% confidence limits. All sample origins were known (rows), but assigned to ocean basins based on the training subset of mitogenomes.

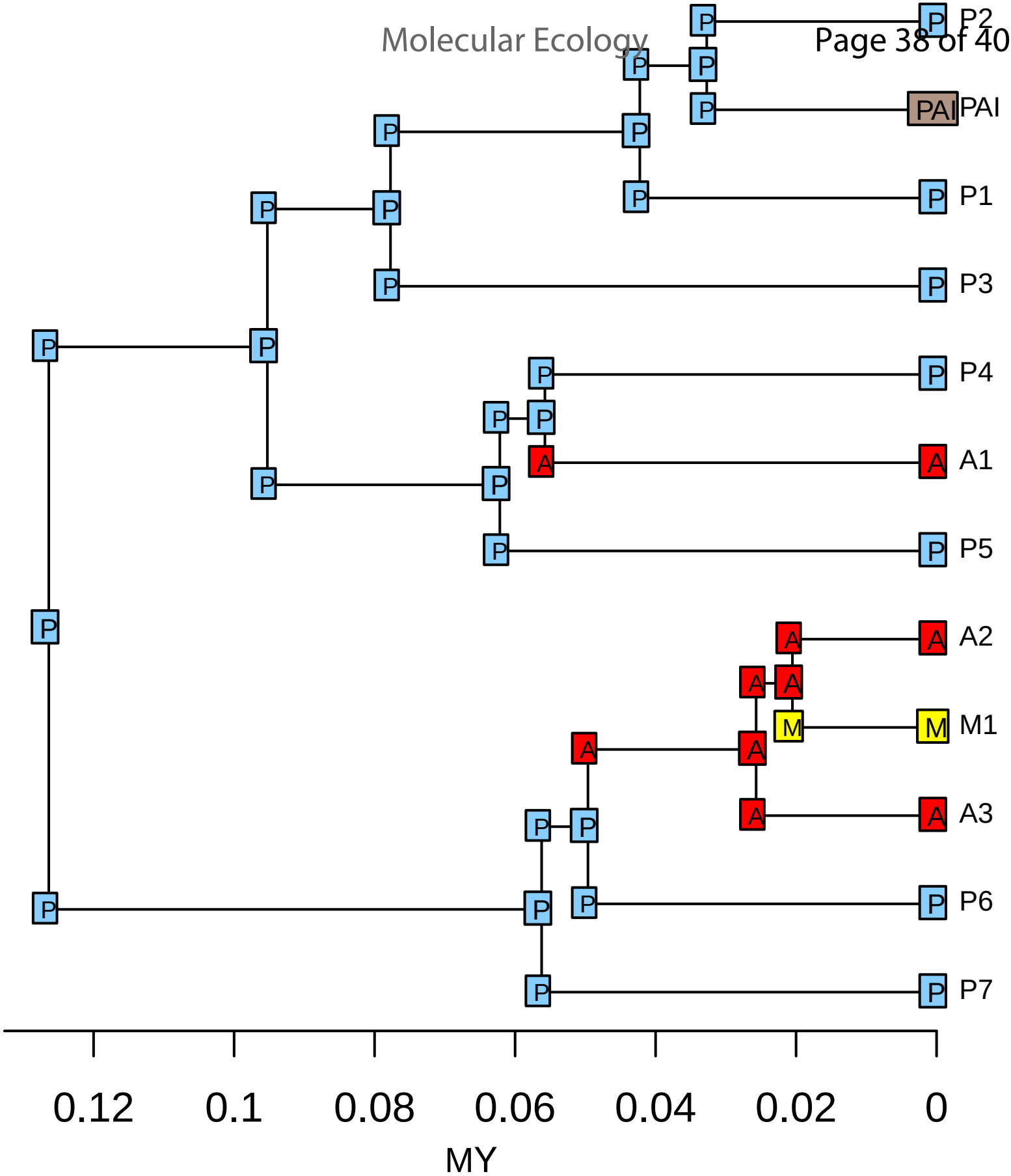
	Atlantic	Pacific	pct correct	LCI	UCI
Atlantic	39	11	78	64.0	88.5
Pacific	0	124	100	97.1	100


New
Published


mt55_Mex
mt47_Japan
mt72_Palmyra
mt24_Mex *
mt07_NPac
mt26_Mex
mt23_Galapagos
mt76_NWAtl
mt32_GoM
mt13_Npac_HI
mt74_Russia
mt81_ENPac
mt52_Tasmania
mt50_Tasmania
mt61_Peru
mt63_ENPac
mt21_NZ
mt71_NZ
mt80_NPac
mt01_NEPac
mt64_Chile
mt66_Chile
mt57_Galapagos
mt06_NPac
mt53_Tasmania
mt33_Maldives
mt04_Tasmania_strand
mt34_NZ
mt03_WatI_NEAtl_Npac
mt45_Philippines
mt78_NZ2
mt78_NZ
mt10_HI_Mex_Chile
mt60_Peru
mt37_NZ
mt54_Tasmania *
mt43_NZ
mt59_Mex
mt41_NZ
mt39_NZ
mt11_Mex_Rus_ChathamIs
mt22_NWAtl_GoM
mt02_NWAtl_GoM_NEAtl *
mt19_NZ
mt69_Chile *
mt29_HI
mt42_NZ
mt70_HongKong
mt18_NZ
mt15_Mex
mt73_NEPac
mt38_NZ
mt31_NZ
mt27_ENPac_Peru
mt48_Npac
mt44_ENPac
mt68_Chile
mt20_NZ_Peru
mt58_ENPac
mt08_Marianas
mt62_Peru
mt05_NWatl_GoM
mt30_GoM
mt79_NWAtl
mt56_NWAtl
mt16_Med
mt75_Med
mt09_NWAtl
mt17_NWAtl
mt77_NWAtl
mt12_Npac
mt51_Tasmania
mt28_Palmyra
mt49_Npac
mt46_GoC
mt35_NZ
mt67_Chile
mt65_Chile
mt40_NZ
mt14_GoC
mt25_Mex
P1
P2
PAI
P3
P4
A1
P5
A2
M1
A3
P6
P7
1.0
0.99
0.90
1.0
1.0
1.0
0.93
1.0
1.0
0.66
0.86
1.0
1.0
1.0
1.0
1.0
1.0
0.125
0.1
0.075
0.05
0.025
0
Million years


10 samples
1 sample
Indian
Atlantic
Pacific
Med.
GoMx


"ice-house" conditions
LGM
Eemian warm period
Effective population size (x10^4)
3.5
3
2.5
2
1.5
1
0.5
0
10^4
10^5
10^6
10^7
Years before present


P2
PAI
P1
P3
P4
A1
P5
A2
M1
A3
P6
P7
0.12
0.1
0.08
0.06
0.04
0.02
0
MY

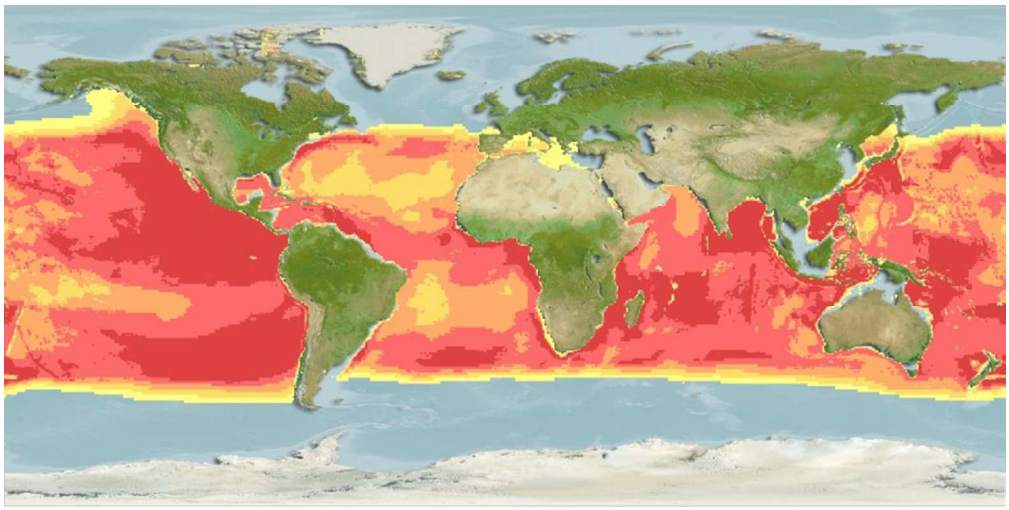
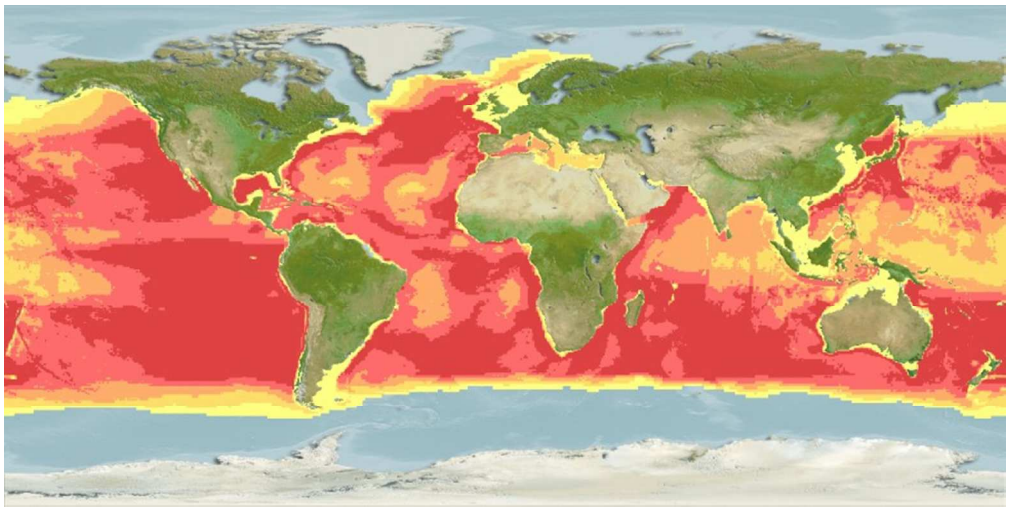


Figure 6. AquaMaps environmental envelope models for distribution of female sperm whale habitat mapped to environmental conditions for a) last glacial maximum and b) current habitat. Dark red color indicates core suitable habitat; yellow indicates marginal habitat.

317x158mm (72 x 72 DPI)



317x158mm (72 x 72 DPI)